DOCUMENT RESUME

ED 142 978 CS 003 616

AUTHOR Hoppe, Richard B.
TITLE Acquisition and Retention of Complex Semantic Ideas. Final Report.
INSTITUTION Kenyon Coll., Gambier, Ohio.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C. Office of Research Grants.
BUREAU NO 4-0404
PUB DATE 9 Jul 77
GRANT NIE-G-74-0010
NOTE 45p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Cognitive Processes; College Students; *Learning Processes; *Memory; *Psychological Studies; *Reading Comprehension; *Reading Research; Recall (Psychological); Retention
IDENTIFIERS *Prose Learning

ABSTRACT

This report describes four experiments on various aspects of memory for connected discourse. The first experiment dealt with preexperimental dispositions to respond to sentences and showed that people do come to experiments with biases; therefore, control of materials is necessary in this research area. The second experiment studied the acquisition process, by inserting recognition probes during acquisition, and showed that specific memory and inferential memory develop at different rates, suggesting that different processes are involved. The third experiment varied the information overlap within acquisition lists; the manipulation had minimal effects upon the acquisition of inferential memory. The final experiment studied retention and showed that the ability to discriminate specific acquisition sentences from inferences is stable over a one-week interval, indicating again the distinction between specific memory and inferential memory. A tentative theoretical scheme that suggests ways of accounting for both specific memory and inferential memory is outlined. (Author/AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED142978

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

FINAL REPORT

PROJECT NO. 4-0404
GRANT NO. NIE-G-74-0010

ACQUISITION AND RETENTION OF COMPLEX SEMANTIC IDEAS

RICHARD B. HOPPE
KENYON COLLEGE
GAMBIER, OHIO 43022

July 9, 1977

The research reported herein was performed pursuant to a grant with the National Institute of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official National Institute of Education position or policy.

U.S. DEPARTMENT OF

HEALTH, EDUCATION AND WELFARE

NATIONAL INSTITUTE OF EDUCATION

OFFICE OF RESEARCH GRANTS

CS003616

ABSTRACT

This report describes four experiments on various aspects of memory for connected discourse. Prior work has shown that people appear to abstract the semantic content of sentences, constructing representations in memory which contain no information about the specific form of the input. In retention tests, people 'remember' information inferrable from several input sentences considered jointly.

The first experiment dealt with pre-experimental dispositions to respond to sentences, and showed that people do come to experiments with biases, and therefore control of materials is necessary in this research area. The second experiment studied the acquisition process by inserting recognition probes during acquisition, and showed that specific memory and inferential memory develop at different rates, suggesting different memorial processes. The third experiment varied the information overlap within acquisition lists. The manipulation had minimal effects upon the acquisition of inferential memory. The final experiment studied retention, and showed that the ability to discriminate specific acquisition sentences from inferences is stable over a one week interval, indicating again the distinction between specific memory and inferential memory. A tentative theoretical scheme which suggests ways of accounting for both specific memory and inferential memory is outlined.

The author gratefully acknowledges the assistance of Constance A. Howes, Sylvia B. Robbins, and Dennis P. Wipper, all of whom participated in various stages of the research reported herein. In addition, I am grateful to Professors Bruce Meadows and Randie Timpe of the Mt. Vernon Nazarene College for their help in obtaining experimental participants.

INTRODUCTION

The purpose of the research reported here was to study the acquisition and retention of complex semantic ideas that are experienced by people as a series of sentences presented nonconsecutively in time. Four separate experiments were performed, bearing respectively upon (a) pre-experimental dispositions to respond to test materials, (b) the acquisition process (two experiments), and (c) retention over extended intervals. Following this general introduction and review, the experiments will be presented individually. Finally, a general discussion section will summarize the research and suggest some conclusions that may be drawn from the data.

The theoretical background and methodological approach used in this research derive from the work of Bransford and Franks (1971), who developed a methodology for studying the acquisition of complex semantic ideas. Their procedure was as follows. They started with what they termed an "idea set", expressed as a complex sentence (e.g., The ants in the kitchen ate the sweet jelly which was on the table, termed a "FOUR" since it contains four basic semantic relations). The semantic relations contained in a FOUR can be partially expressed in less complex sentences (e.g., The ants in the kitchen ate the sweet jelly, a THREE; The ants in the kitchen ate the jelly, a TWO; The ants were in the kitchen, a ONE). For a given FOUR there are 3 THREEs, 4 TWOs, and 4 ONEs, or a total of 12 sentences comprising the idea set. A subset of these THREEs, TWOs, and ONEs, six from each of four idea sets, were presented to Ss in an acquisition task which required Ss to both hold the sentence in short-term memory for a few seconds and to attend to the semantic aspects (the meaning) of the acquisition sentences. The FOURs were not presented in acquisition, nor were 1 THREE, 2 TWOs, and 2 ONEs. Subjects were not informed that a test phase would follow the acquisition phase. Approximately five minutes after the completion of the acquisition phase, Ss were read another series of sentences, and were instructed to answer whether or not each test sentence had or had not occurred in the acquisition phase, and to rate their confidence in their answer. In various experiments reported in Bransford and Franks (1971), the test sentences included (a) old ONEs, TWOs, and THREEs, which had been presented in acquisition; (b) new ONEs, TWOs, and THREEs, which had not been previously presented; (c) the FOURs, which had not been presented in acquisition; and (d) NONCASEs, complex centences of the form of FOURs but which were formed by either changing the semantic relations of one of the FOURs (e.g., The dog chased the cat versus The cat chased the dog) or by combining elements from two FOURs. The NONCASEs, of course, had not occurred in acquisition.

The three important results of Bransford and Franks' work are (a) FOURs, which were not presented in acquisition, but which do include all the semantic relations of the complex idea set, were 'recognized' with the highest confidence; (b) NONCASEs, which were not presented in acquisition and which cross over between idea sets, were rejected with high confidence; and (c) Ss' recognition confidence was a monotonically increasing function of inclusiveness (degree to which a sentence exhausts the content of the FOUR from which it was derived). Bransford and Franks' results showed the Ss cannot distinguish between (a) and (b); that is, they cannot discriminate derivative sentences presented in acquisition from those not presented. Rather, the recognition responses depended upon the degree to which a given test sentence exhausted the content of the FOUR from which it was derived. Thus, THREEs were 'recognized' with greater confidence than TWOs, and TWOs with greater confidence than ONEs, irrespective of whether a given sentence had in fact occurred in

acquisition. The FOURs, which had not occurred in acquisition, were nonetheless most confidently 'recognized' as having been presented. NONCASEs, on the other hand, were confidently rejected as not having been presented. Bransford and Franks interpret these findings to mean that in acquisition Ss were "integrating information across various acquisition sentences experienced nonconsecutively in time [1971, p. 348]." In the test phase, Ss were presumably performing a matching task, comparing the test sentence with the representation constructed in acquisition. If a given sentence was congruent with the representation, S answered "Yes", he had seen it before. To the degree that the test sentence exhausted the content of the representation, the Ss were more confident of their responses.

The basic phenomenon of a monotonic relationship between test sentence inclusiveness and Ss' confidence ratings has been replicated a number of times (Cofer, 1973; Franks & Bransford, 1972; Gould & Bourland, Note 1; Katz, 1973). Essentially, all of these results are implying that in acquisition Ss do not simply store a representation of each input sentence as it occurs, but rather are actively constructing semantic wholes out of the parts presented in the various acquisition sentences. (These data do not speak clearly to the question of what is going on in the test phase; the brief account in the paragraph preceding this one seems to be an implicit working assumption in all the cited studies.) Taken with other work like that of Sachs (1967) and Jarvella (1971), as well as some of the work on figural memory (Franks & Bransford, 1971; Strange, Keeney, Kessel, & Jenkins, 1970), this work indicates that the reconstruction approach, using recognition-memory test tasks, can yield good data bearing on the problem of representation in memory.

There is, however, at least one major problem with the Bransford-Franks methodology as it has been used to study memory. In order to interpret the form of the recognition curves for FOURs, THREEs, TWOs, ONEs, and NONCASEs as being due to processes occurring during and after the acquisition phase of the experiment, it must be assumed that the experimental test materials are pre-experimentally neutral. That is, one can consider the test task, where S is to choose which of a series of sentences occurred in acquisition, as essentially a stimulus discrimination task in which the relevant dimension for the discriminations is what might be called 'familiarity', or more precisely, 'familiarity in the experimental context'. In order to attribute Ss' behavior with respect to the test sentences to acquisition processes, it is necessary to assume that the test sentences are homogenous with respect to any other stimulus dimensions; that the sentences used in the test phase differ only along the 'familiarity in the experimental context' dimension. Several years ago my students and I looked at this aspect of the Bransford-Franks procedure, and found that the test sentences used by Bransford and Franks (1971) differ systematically along a dimension we called pre-experimental oddity (Gould & Bourland, Note 1). That is, using the Bransford and Franks test FOURs and NONCASEs, we found that naive Ss could reliably sort them into two categories -- "odd" and "normal" -- without any acquisition experience. By varying the relationship of FOURs to NONCASEs with respect to oddity, we both lowered the confidence with which FOURs are 'recognized' and lowered the confidence with which NONCASEs are rejected, indicating that this test material stimulus dimension is confounded with the 'familiarity in the experimental context' dimension assumed by Bransford and Franks to be solely responsible for the form of their results.

The first study reported here systematically examines the pre-experimental tendency of subjects to respond differentially to retention test materials in

the absence of an explicit experimental acquisition experience. It does not seem, *a priori*, to be likely that people come to experiments which utilize sentential experimental materials disposed to treat those sentences alike. If the long history of verbal learning in the associationist tradition has taught us nothing else, it shows that people do not treat verbal materials indifferently; even Ebbinghaus' (1885) nonsense syllables were not equivalent with respect to important variables.

The necessity to know the Ss' pre-experimental response dispositions becomes more pressing in light of the theoretical framework within which Bransford and Franks' methodology was developed. The theory of memory held is outlined in Barclay (1973) and in Bransford, Barclay, and Franks (1972). Briefly, they hold that what Ss are doing in acquisition is not simply decoding the incoming sentences, deriving from each a deep structure which specified a complex of semantic features which is the 'meaning' of the sentence, but rather that Ss are actively constructing a semantic representation from the incoming sentences. This semantic representation is derived from "the interplay of sentential information, the context of knowledge to which that information is assimilated, task demands, and the assimilative processes themselves, including integrative and logical operations. These semantic representations in turn serve as memory representations [Barclay, 1973, p. 232]." In other words, the incoming information interacts with pre-experimental knowledge to produce the memory representation which, as a consequence, contains more than was contained in the experimental acquisition materials. To this point, however, none of the research has been directed at ascertaining the relative roles of acquisition materials versus pre-experimental knowledge, and the first part of the proposed research is directed at that question.

The research discussed to this point has been mainly directed at first establishing the phenomenon's reliability and then at ascertaining some of the variables affecting the phenomenon. Katz (1973) varied an instructional variable, and found that the form of the recognition curves varies as a function of the degree to which Ss' attention is directed at the semantic dimension (meaning) of the test sentences. When Ss are directed to attend to the meaning of the test sentences, the monotonic relationship between recognition confidence and sentence inclusiveness disappears, though the FOURs are still recognized with the greatest confidence.

Cofer (1973) has looked briefly at the effects of a filled retention interval on Ss' 'retention' responses. He compared the confidence with which Ss 'recognized' the various types (FOURs, THREEs, TWOs, and ONEs) of test sentence as a function of their position in the test list, and found that Ss 'recognition' confidence for FOURs increased as the FOUR occurred later in the test series, while the confidence for THREEs, TWOs, and ONEs decreased over test list position. This, however, is not purely a retention test, though, since in addition to the retention interval increasing over test series position, Ss' exposure to related sentences increases. That is, given that the target sentence is in the last position of the test series versus the first position, when the sentence is in the last position it is not only further in time from the acquisition series, but it is also preceded by a long series of sentences related to those S is to recognize. This seems not to be a simple test of retention over varying intervals so much as it is a test of the effects of interfering interpolated tasks on retention.

The second two experiments reported deal with the acquisition process itself. The first experiment varied the amount of acquisition which preceded the retention test, and the second experiment looked at the effects of supplying information in acquisition sentences which overlapped from acquisition sentence to acquisition sentence. The effort here is to begin to look at the details of the phenomenon found by Bransford and Franks (1971), and especially, in the first experiment, to study the development of people's inability to discriminate new sentences from old sentences.

The fourth experiment reported, the retention study, looks at the long-term (one week) retention of the semantic wholes presumably formed during the standard 24-sentence Bransford-Franks acquisition procedure. Three retention intervals (one hour, one day, and one week) were interposed between acquisition and test. The question here was whether the phenomenon reported by Bransford and Franks (1971) and others was an enduring one, and therefore of both theoretical and applied interest, or of relatively transient nature, and therefore of principally theoretical interest.

THE EXPERIMENTS

Experiment 1: Pre-Experimental Dispositions to Respond.

As noted above, the interpretation of the phenomenon reported by Bransford and Franks (1971) and others which holds that subjects are constructing a representation of the whole idea which is presented in bits and pieces during acquisition depends in part on the assumption that the test sentences used are pre-experimentally neutral, that there is no tendency to respond differentially to the sentences in the absence of an acquisition experience. That is, in order to attribute the form of the obtained recognition curve to memory for the results of acquisition processing, one must assume that the curve would have been flat had the subject not had an acquisition experience. Put another way, one must assume that there are no effective stimulus dimensions in the test materials except what might be called 'familiarity in the experimental context'. The assumption implies the need for a control condition which tests for the possibility that human subjects bring to an experiment considerable experience with language and sentences, which experience may have some effect upon their behavior in the experimental situation, interacting with the nominal experimental manipulations. If subjects do bring pre-experimental biases to the experiment, then Bransford and Franks' results are a product of the interaction between the pre-experimental biases and acquisition processes, and they cannot then be interpreted as the product of acquisition alone. That pre-experimental variables affect subjects' memory for verbal stimuli is well documented in the traditional verbal learning literature. Variables like frequency of occurrence in the language, inter-item associative strength, or imagery are almost automatically taken into account when one is designing materials for an experiment using words. It seems reasonable to suppose that subjects do not come to a memory experiment which employs sentences without some pre-experimentally acquired experience that may lead them to respond differentially to the test sentences in the absence of an experimental acquisition experience, and that therefore some sort of 'no-acquisition' control condition might be necessary. Experiment 1 provides such a control condition.

Method

Subjects. A total of 231 subjects served in Experiment 1. They were volunteers drawn from introductory psychology courses at Kenyon College. Of the total, 90 served in the norming of materials, and 141 served in the experiment itself.

Materials. Four complete sets of materials were prepared, one set for each of the materials conditions. Each set consisted of (a) four complex sentences, each containing four simple ideas which together define a FOUR (e.g., The young girl in the meadow was picking white flowers); (b) six NONCASEs, formed by extracting simple ideas from two or more FOURs and combining them in sentences similar in form to the FOURs; and (c) ONEs, TWOs, and THREEs, formed by extracting from each FOUR one, two, or three of the simple ideas and constructing less complex sentences from them. From each FOUR were derived four ONEs, four TWOs, and three THREES. Thus the total materials for each of the four materials conditions was four FOURs, 12 THREEs, 16 TWOs, 16 ONEs, and six NONCASEs.

The FOURs and NONCASEs used in the experiment were selected from a larger set according to their relative oddity as determined by norming procedures. Norming was accomplished by requiring 90 subjects not otherwise involved in the experimentation to rate the FOURs and NONCASEs, presented to them in the same manner as in the main experiment, for relative oddity. These subjects were instructed to rate each sentence according to how odd it seemed to them on a scale of 1 to 7, with 7 indicating high oddity. Each set of materials was normed separately, with the four norming groups ranging in size from 18 to 26 subjects. The norming procedure yielded four sets of materials with the four possible relative oddity relations between FOURs and NONCASES: Normal FOURs/Odd NONCASEs (N-O), Odd FOURs/Normal NONCASEs (O-N), Normal FOURs/Normal NONCASEs (N-N), and Odd FOURs/Odd NONCASEs (O-O). Only the FOURs and NONCASEs were normed; the THREEs, TWOs and ONEs were derived from the FOURs selected by the norming procedure, and were not themselves normed. The median oddity ratings for the four sets of materials are shown in Table 1.

Insert Table 1 about here

The acquisition list for the standard Bransford and Franks condition was constructed as follows. The acquisition list was composed of 24 sentences, two THREEs, two TWOs, and two ONEs from each of the four FOURs of the particular oddity condition materials. Each block of four sentences on the acquisition list contained one representative derived from each of the FOURs, and the order of presentation was randomized within these blocks with the constraint that two sentences derived from the same FOUR could not appear in sequence. The acquisition sentences derived from each FOUR exhausted the semantic content of that FOUR within the full acquisition set. The order of presentation to one subgroup within each procedure by materials cell was 1-24; to the other subgroup it was 24-1.

The test list used in all three procedural conditions was constructed in a similar manner. It was composed of 30 sentences of the following types: (a) the two ONEs, two TWOs, and one THREE from each set that were not presented in acquisition; (b) the four FOURs, not presented in acquisition; and (c) the six NONCASEs, also not presented in acquisition. The test list was randomized in the same manner as the acquisition list, with the addition of one NONCASE to each block of four sentences. The order of presentation to one subgroup within each procedure by materials cell was 1-30; to the other subgroup it was 30-1.

Procedure. Three procedural conditions were crossed with the four materials conditions to produce 12 combinations of conditions. The three procedural conditions were two 'no-acquisition' conditions and one standard Bransford and Franks acquisition condition. The two 'no-acquisition' conditions differed in their instructions. The 'soft' instructions directed subjects to rate the "likelihood that you could have heard each sentence sometime in the past... Remember, the question is not "Have you heard the sentence?', but rather, Is it likely that you could have heard it?" The 'hard' instructions were exact replicates of the recognition instructions used in the standard acquisition group, and directed subjects to answer whether or not they had heard the sentence before, and rate the certainty. In the acquisition condition, subjects in the acquisition phase heard 24 THREEs, TWOs, and ONEs, drawn from the set of materials appropriate to the materials condition. After each sentence was read, the experimenter read four numbers, which the subjects repeated in unison. Then an elliptical question about the preceding sentence was read, and subjects answered on an answer sheet. A four to five minute break separated the acquisition and test phases. In the test phase, thirty test sentences were read to the subject. Following each, the subjects were to answer whether or not they had heard the sentence before, and rate their confidence with respect to the answer on a scale of five points, "1" indicating low confidence. Subjects in the acquisition condition were unaware that a test phase was involved until it occurred.

Subjects were unsystematically assignedto the twelve cells of the design. Since no statistical tests involving the homogeneity of variance assumption of the analysis of variance were contemplated, no particular effort to maintain group equality was made, and the only constraint on group size was that each be large enough to ensure reasonable stability of the results. Group sizes ranged from 8 to 11 subjects in the no-acquisition, soft instructions condition, and from 12 to 14 in the other two procedural conditions. Within each cell, two subgroups were run to counterbalance the order of presentation of sentences. All materials were tape-recorded for presentation.

Results

The first assumption tested, that subjects come to an experiment involving sentential materials with some predispositions to respond differentially to the test materials, receives support in the data. The results of the two no-acquisition conditions for the four materials conditions are shown on Figure 1. (The numbers shown on this and the other figures in this paper were

---

Insert Figure 1 about here

---

derived by assigning a plus sign to each confidence rating for which the subject answered that he/she had heard the sentence, and a negative sign to those for which the answer was that he/she had not heard the sentence. The data points are means over subjects and individual sentences within types.) The most striking result in the figure is the pattern displayed by the N-O sentences, those comparable to the materials used by Bransford and Franks (1971) and other researchers. Under both 'hard' and 'soft' instructions subjects respond systematically differently to the various types of sentences. Of the four materials conditions, the N-O is most contaminated by pre-experimental response biases, and as a consequence

data resulting from this sort of condition in the Bransford-Franks procedure are the least dependable with respect to conclusions about what is happening in the acquisition phase. The least contaminated materials appear to be the O-O sentences, those in which FOURs and NONCASEs are roughly equal in pre-experimental oddity. In all four materials conditions the NONCASEs are rejected with strong confidence, though not as strong as in prior work involving an acquisition phase nor as confidently as in the acquisition condition in the present experiment. (Only the N-O and O-O conditions appear on the figure.)

The results of the acquisition condition for the four materials conditions are shown in Figure 2. Note that for the most part these data replicate those

---

Insert Figure 2 about here

---

previously reported in the literature, with the FOURs rated higher than the THREEs, the THREEs higher than the TWOs, and so on. The only failure to replicate occurs in the O-N condition, where the FOURs are rated just lower than the THREEs. In all four materials conditions the NONCASEs are rejected with great confidence, although the amount of change in the NONCASEs' rating from the no-acquisition conditions is relatively small, amounting to only about one scale unit. In terms of change from no-acquisition, the greatest effect is on the N-N FOURs and the N-O ONEs, with the FOURs being raised over five scale units from the 'hard' instructions no-acquisition condition to the acquisition condition. The general effect of the acquisition experience was to raise the FOURs, THREEs, and TWOs, and to lower the ONEs and NONCASEs, as compared with the 'hard' no-acquisition condition. Recall that these two procedural conditions had identical instructions in the test phase; they differed only in that in the no-acquisition condition subjects had no prior experimental experience with sentences related to the test sentences.

Discussion. The data of Experiment 1 clearly indicate that the assumption discussed in the introduction is invalid, that subjects do come to experiments with pre-experimental dispositions to respond to the test materials independent of any explicit experimental acquisition experience. Of the four materials conditions tested, the O-O seems most neutral, especially with respect to the FOURs, THREEs, TWOs, and ONEs, which deviate only slightly from zero in the 'hard' instructional condition. As a consequence, interpretation of retention tests involving the other sorts of sentences is rendered problematic in the absence of such a 'no-acquisition' control, while those involving sentences like the O-O sentences are less problematic. As a result, in the subsequent experimentation reported in this document, only the O-O materials are used.

Experiment 2: Formation of Semantic Wholes in Acquisition.

The purpose of Experiment 2 was to look at the construction of the hypothesized 'semantic wholes' during the acquisition phase. If the semantic wholes are constructed in the course of the acquisition phase, then a test probe inserted early in acquisition should produce results that are clearly different from those produced at the end of acquisition. In order to test this assertion, it is necessary to control for pre-experimental dispositions to respond, and therefore the O-O materials from Experiment 1, those materials least contaminated by

pre-experimental biases, were used in this experiment. In addition, it was necessary in Experiment 2 to control the density of information within the acquisition list. In previous research, including Experiment 1 of this paper, the only constraint on the acquisition list was that over the whole list all the simple ideas of each FOUR be presented. In the present experiment, this constraint was applied to thirds of the list, so that probes could be inserted after each third of acquisition.

Method

Subjects. Thirty-six volunteers from introductory psychology classes at the Mount Vernon Nazarene College served as subjects.

Materials. The O-O materials of Experiment 1 were used as experimental materials in Experiment 2. The acquisition list was constructed in the same manner as in Experiment 1 except that within each third of the list (eight sentences) the two representatives from each FOUR exhausted the content of the FOUR. Within the first third of the list, two of the FOURs were represented by a THREE and a ONE, and the other two FOURs were represented by two TWOs. In the second third of the list, this representation scheme was reversed. In the final third of the acquisition list, all the FOURs were represented by THREE-ONE pairs. The precise sentences composing the acquisition list changed from Experiment 1 so as to allow insertion of a few acquisition sentences into the test sentence list. The test list was constructed in the same manner as in Experiment 1 with the exception that five of the sentences from the first third of the acquisition list, one THREE, three TWOs, and one ONE, were inserted into the test series. In addition, another ONE and another TWO from the second third of the acquisition list were inserted into the test series. Thus, for those subjects who heard only the first eight sentences of the acquisition list, five of the 30 test sentences were old, and for those subjects who heard the first two-thirds of the acquisition list, seven of the 30 test sentences were old.

Procedure. Twelve subjects were assigned randomly to each of three acquisition conditions: 1/3, in which subjects heard only the first eight acquisition sentences before test; 2/3, in which subjects heard the first sixteen acquisition sentences before test; and 3/3, in which subjects heard all 24 acquisition sentences before test. The last condition replicates the standard Bransford and Franks (1971) procedure. Within each acquisition condition, half the subjects heard the acquisition and test sentences in one order, and the other half heard the sentences in reverse order.

Instructions to subjects in all three conditions were identical to those used in the acquisition condition of Experiment 1. Once again, subjects were not aware prior to its occurrence that a test phase would occur.

Results

The results of Experiment 2 will be considered in conjunction with the results of the no-acquisition, 'hard' instructions condition of Experiment 1. The latter condition used the same instructions as did the former in the test phase, and as a consequence the no-acquisition data form the starting point for the construction of the semantic wholes which is probed in Experiment 2.

Figure 3 shows the data for the three probe conditions along with the

---

Insert Figure 3 about here

---

the no-acquisition data of the 0-0 condition from Experiment 1. The curves for new sentences show a clear progression from the no-acquisition condition through the 1/3, 2/3, and 3/3 conditions. With a little acquisition experience, subjects move from essentially no tendency to systematically 'recognize' the new sentences in the no-acquisition condition to a mild tendency to reject those sentences. With increasing experience in acquisition, the curve moves upward, and begins to take on the characteristic form. With the full 24-sentence acquisition experience, the curve very nearly replicates the data of all prior research.

The old sentences, those that appeared in both the acquisition and Test phases of Experiment 2, show a quite different pattern across the levels of acquisition experience. The data are shown on Table 2. With only eight acquisition sentences, the old sentences in test move directly up to fairly high positive values, and maintain virtually the same level through the increasing amounts of acquisition.

Insert Table 2 about here

The results of Experiment 2 were analyzed using two-way analyses of variance with repeated measures on both variables (old vs. new, and sentence type). Each acquisition condition was analyzed separately; i.e., three separate analyses of variance were performed, and only the THREEs, TWOs, and ONEs were included in the analyses, since it is in these three sentence types that the old/new comparison is possible. Three separate analyses of variance were performed in order that further determinations of the variance components and their changes across acquisition conditions might be made. These latter analyses are reported in the joint results section. Table 3 displays the outcome of the

---

Insert Table 3 about here

ANOVAs. As may be seen, the old/new variable is the only significant effect in the 1/3 condition, while in the 2/3 conditions both main effects are significant and in the 3/3 conditions both main effects and their interaction are statistically significant. In other words, throughout acquisition subjects are reliably discriminating old from new sentences, while in the later stages of acquisition they begin to differentiate among the sentence types. (Recall again that these analyses involve only THREEs, TWOs, and ONEs.)

Discussion. The principal conclusions that emerge from the data of Experiment 2 are that while after a full 24-sentence acquisition experience the curves for old and new sentences are similar in form and location, the route by which the curves get to their final position is different over the course of the acquisition process. Moreover, even after a full 24-sentence acquisition phase, there is still a statistically significant difference between old and new sentences; subjects can still discriminate that which they heard from that which is merely consistent with what they heard in acquisition. In the joint results

section these data will be brought into direct juxtaposition with those from the retention study (Experiment 4), in order to trace the course of acquisition and retention of the semantic wholes.

Experiment 3: Information Overlap in Acquisition.

As discussed in the general introduction, the explanation offered by Bransford and Franks (1971) and others (e.g., Bransford, Barclay & Franks, 1972) asserts that the phenomena observed to occur in the recognition tests (failure to discriminate old sentences from new, and responding as a function of sentence 'exhaustiveness') are a consequence of processes supposed to occur during acquisition. Specifically, it is suggested that during acquisition people are not simply storing a representation of the input sentences in memory, but are rather abstracting from the sentences their semantic content, and constructing a representation of the whole content out of parts contained in the several input sentences. In this portion of the research, the materials were constructed so that within each third of the acquisition list, two of the idea sets (FOURs) was represented as was the case in the preceding experiment (Experiment 2). The other two idea sets, however, were represented by two sentences which contained a common semantic relation. If, as Bransford, et al. (1972) suggest, the input sentences are not treated as separate linguistic objects, but serve as part of the raw material for the construction of wholistic semantic relations, then this manipulation should facilitate their construction as indexed by probe tests inserted in a manner similar to that of Experiment 2.

Method

Subjects. Thirty-six introductory psychology students at Kenyon College served as subjects. All were volunteers.

Materials. The O-O materials of Experiment 1 were used as experimental materials in Experiment 3. The acquisition list was constructed in a slightly different manner, however. Two of the idea sets were randomly selected to be 'overlap' sets, and the other two were designated 'non-overlap' sets. The acquisition sentences for the non-overlap sets were selected according to the same constraints used in Experiment 2. The acquisition sentences from the overlap sets were selected such that, within each of the first two thirds of the acquisition list, each was represented by a THREE-TWO pair. As a consequence, one semantic relation was repeated within each of the first two thirds of the acquisition list, since the THREE-TWO pair was selected to exhaust the content of the idea set. The final third was constructed as in Experiment 2. The recognition test list was identical to that of Experiment 2.

Procedure. The procedure of Experiment 3 was identical to that of Experiment 2.

Results. Since the principal evidence adduced to support the explanation involving semantic wholes is subjects' willingness to accept new sentences, the analyses in this experiment were performed upon only the new sentences in the test phase. Figure 4 shows the new test data for the overlap sentences -- those new sentences derived from idea sets presented in overlapping acquisition sentences. Over the three acquisition conditions the curve moves upward in a

---

Insert Figure 4 about here

---

manner similar to that shown in the previous experiment, although the initial rejection of FOURs that appeared in Experiment 2 does not appear here. The curves for the nonoverlap test sentences are shown on Figure 5. Once again,

---

Insert Figure 5 about here

---

there is a progression, with one major difference: In the 3/3 acquisition condition, the test FOURs in the non-overlap condition are well below the THREEs. This is the first substantial failure to replicate the monotonic relationship between confidence rating and test sentence exhaustiveness (using standard recognition instructions) reported in the literature. The overlap variable had little effect upon acquisition of the semantic wholes. Analyses of variance, conducted in the same manner as in Experiment 2, showed that the overlap variable was effective only in the 2/3 acquisition condition, where overlap sentences were recognized with somewhat greater confidence than non-overlap sentences. Table 4 shows the outcomes of the ANOVAs. The significant interaction in the

---

Insert Table 4 about here

2/3 condition was due principally to the wide separation between overlap and non-overlap FOURs, the overlap FOURs being accepted with relatively high confidence. This tendency to accept overlap FOURs with higher confidence persists in all three acquisition conditions, though not at levels statistically significant in the 1/3 and 3/3 conditions.

Discussion. The results suggest that there is some effect of overlapping information in acquisition upon later test performance, but that this is a relatively weak variable. The effect is seen principally in the middle stages of acquisition, with the effect being diminished with more acquisition experience. The effect may be in part explained by the acquisition material arrangement necessary to achieve the overlap manipulation. The 1/3 and 2/3 acquisition sentences in the overlap condition were all THREEs and TWOs. These subjects did not hear ONEs in acquisition, and therefore may have been biased toward longer sentences in recognition. It is of note in this connection that the 3/3 subjects, who heard ONEs, did not give FOURs as high a rating as did the 1/3 and 2/3 subjects. This result is consistent with Flagg (1976).

Experiment 4: Retention of Semantic Wholes.

The final experiment in the series was a retention study. All prior work with the phenomenon used a four to five minute interval between the end of acquisition and the beginning of the recognition test. This experiment used three intervals: one hour, one day (24 hours) and one week (168 hours). The purpose of the experiment is to ascertain the durability of the two basic phenomena in the Bransford-Franks paradigm: (a) the inability, or near inability, to distinguish old test sentences from new sentences, and (b) the display of a monotonic relationship between test sentence exhaustiveness and confidence ratings.

Method

Subjects. Thirty-six introductory psychology students at Kenyon College served as subjects. All were volunteers.

Materials. The materials used were those of Experiment 2.

Procedure. Twelve subjects were assigned randomly to each of the three retention interval groups. Within each group, two subgroups were randomly formed to allow counterbalancing of the order of presentation of materials. Each group heard all 24 acquisition sentences, and, after the appropriate interval, 30 test sentences which contained old and new THREEs, TWOs, and ONEs, the FOURs, and six NONCASEs, the same test materials as in Experiment 2. Subjects were not informed that a retention test would occur.

Results. The results were analyzed in a manner similar to that used in Experiment 2. The results for new test sentences are shown on Figure 6. The

---

Insert Figure 6 about here

---

curves for each of the retention conditions are very similar, with a slight upward drift over retention intervals. Table 5 shows the data for the old

---

Insert Table 5 about here

THREEs, TWOs, and ONEs. The data are presented in a table because the data points are too close together to be clearly represented on a figure, which is itself a striking result which will be discussed later.

Three analyses of variance were performed, a two-way ANOVA with repeated measures on both factors for each retention condition. Once again, the separate analyses were performed to enable later subanalyses to be discussed in the next section. The outcomes of the ANOVAs are shown on Table 6. As may be seen,

---

Insert Table 6 about here

---

both main effects, sentence type and old/new, were statistically significant in all three retention conditions, while the interaction of the two was significant in the 1-Day and 1-Week conditions.

Discussion. The remarkable outcome of Experiment 4 is that subjects were able to significantly discriminate old from new test sentences over all three retention intervals; i.e., the phenomenon which first appeared in Experiment 2 replicates over the three retention conditions of the present experiment. These

data demonstrate that the phenomenon described by Bransford and Franks (1971) and Curnow (1969) is not replicable under conditions where the pre-experimental biases of subjects are controlled by the use of pre-experimentally neutral materials.

Joint Results.

This section brings together selected results from three of the experiments, Experiments 1, 2, and 4, in order to illustrate the development and retention of the semantic wholes from no acquisition through full acquisition with a 1-Week retention interval. Figure 7 shows the progress of new FOURs, THREEs, TWOs, and ONEs from no acquisition to 1 week retention. All four types of

---

Insert Figure 7 about here

---

sentences follow similar courses across acquisition and rentention conditions. They all start around zero -- no tendency to respond differentially. Over the early acquisition period, they all decrease, meaning that subjects are tentatively (and correctly) rejecting them in test. In mid-acquisition, THREEs and TWOs remain steady, while FOURs increase sharply and ONEs continue their decline. With full acquisition, the THREEs and TWOs increase, as do the ONEs, though the latter are still fairly confidently rejected. All four types drop slightly at the 1-Hour retention interval, and then rise again with a 1-Day retention interval.

The picture for the old THREEs, TWOs, and ONEs is quite different. Figure 8 shows the relevant means. Here the old THREEs and TWOs show immediate

---

Insert Figure 8 about here

---

increase as a result of the first eight acquisition sentences, while the ONEs hold steady around zero. Over the acquisition conditions the THREEs and TWOs remain at relatively high levels, while the ONEs increase from the 2/3 to 3/3 acquisition conditions. Over the retention conditions the major change is in the old THREEs, which increase again to very high levels, approaching the ceiling of possible confidence ratings. All three types of old sentences consistently are rated higher than their new counterparts with the exception of a virtual tie between old and new TWOs in the 1-Day retention condition. This consistency is reflected in the statistically significant difference between old and new sentences in all the analyses of variance reported.

A second analysis of the data reported for Experiments 2 and 4 was performed. This analysis estimated the variance attributable to the various main effects and interactions found to be statistically significant in the three experiments. This estimation followed the procedures of Vaughan and Corballis (1969), and produces for each significant effect an estimate of the proportion of the total variance within the experiment that may be attributed to that independent variable or interaction of variables. Table 7 shows these estimates for the three acquisition conditions of Experiment 2 and the three

---

Insert Table 7 about here

---

retention conditions of Experiment 4, together with the estimate of the variance within each experiment due to error. The patterns in variance accounted for across the various conditions show interesting aspects. First, there is the general decrease, with a final rise, in the error variance. This is to be expected, since with increasing acquisition there is increasing control by the independent variables over subjects' behavior, and as a consequence variance due to random factors should decrease. The decrease then simply means that subjects' behavior is becoming more uniform over the conditions, though with a one week interval between acquisition and test the error begins to increase again. With respect to the variance proportions, initially the only factor of significance is the old/new contrast, accounting for 23.2% of the variance in the 1/3 condition. This proportion rises slightly in the 2/3 condition (28.6%), falls sharply in the 3/3 condition (8.6%), then rises again in the 1-Hour and 1-Day conditions, only to fall once again in the 1-Week condition. Meanwhile, variance attributable to the sentence type factor is initially zero, rising to 18.8%, falling slightly to 12.2%, then rising sharply to over 30%, remaining there for the three retention conditions. In the retention conditions, the sentence type variable is a fairly powerful one, accounting for over one-third of the total variance in the respective experimental treatments.

It is of some note that the interaction of the two main effects is of negligible magnitude, accounting for less than 6.0% of the total variance in those conditions in which it was a statistically significant factor. Overall, the sentence type variable is the most powerful, in terms of variance accounted for, while the old/new contrast accounts for a substantial proportion of variance in the early stages of acquisition and in the 1-Hour retention condition. It is also of note that in the 3/3 condition, which replicates the experimental procedure of virtually all prior research using this methodology, all the variables are of relatively little weight with respect to variance accounted for, even though in this condition the error variance was below the average.

GENERAL DISCUSSION

First, to summarize the results of the experimentation: Experiment 1 showed that people do have pre-experimental dispositions to respond differentially to the test sentences even under 'hard' instructions like those used with subjects who have an acquisition experience. The results of Experiment 2 showed that old and new test sentences are discriminable if the pre-experimental dispositions are controlled, and that old and new sentences follow different paths over the course of acquisition, suggesting that similar-appearing curves may be the product of different memorial processes or structures. Experiment 3 suggested that the overlap manipulation in acquisition may have some relatively small effects during acquisition, though the effects are not statistically significant after the standard 24-sentence acquisition set. Experiment 4 showed that both the tendency of subjects to respond to test sentences as a function of the inclusiveness of the sentences and their ability to reliably discriminate old from new sentences persist over a retention interval of one week, suggesting that the phenomena are of some significance for both theoretical and applied purposes.

The theoretical account originally offered for the data reported by Bransford and Franks (1971; Barclay, 1973; Bransford, Barclay, & Franks, 1972) is inadequate to account for the results reported here. The 'semantic wholes' hypothesized in that account provide no means by which people may distinguish old from consistent new sentences, but people can do it. Rather than completely discarding particulars of the input experience, something remains of that experience to influence later recognition choices. An adequate theoretical account must account for both the ability of people to discriminate between old and new sentences, and for their willingness to 'recognize' new sentences which they have not heard in acquisition but which are consistent with or derivable from sentences which they did hear, even though that recognition is not as confident as the recognition of old sentences. That is, the structure of what is learned in acquisition must be described in such a way as to include information regarding the particular occurrence of parts of the whole complex idea set, and also have a structural form that permits the inference of additional relationships among the elements of the idea set. The theory must make certain assumptions about the nature of the permanent lexical memory structure which underlies the ability to comprehend sentence elements, as well as assumptions about a grammatical theory consistent with the lexical assumptions. The grammatical theory must provide a mechanism for understanding the sentence as a whole, as well as a means by which the perceiver can recognize that elements of one sentence are those which appear in another related sentence. The present report will briefly sketch all three components of the theoretical account. It is emphasized that this is as yet a sketch; much remains to be worked out.

The theory of lexical memory in the present account is very similar to several proposed in the literature, and suggests that words are represented as a complex of features, where the features are of at least three general sorts: perceptual/identification, response generation, and semantic. (This general sort of feature theory is found in Gibson, 1971; Smith, Shoben & Rips, 1974; and Wickens, 1970; among others.) The present theory assumes that the permanent representation of a word in lexical memory contains information of three sorts: perceptual/identification information, response generation information, and semantic information, and that the three parts are semi-independent. That is, there is stored at some organizational location information that enables access to the location from the perceptual processing system, information that allows the word to be generated as a response, either spoken or written, and information concerning the meaning (and syntactic class(es)) of the word. While there is neither the space nor the necessity to develop in detail the perceptual/identification and response generation components in this report, a few things should be noted. It is <u>not</u> necessarily assumed that a specific articulatory motor program for a word be stored at the node in lexical memory, but rather it is more likely that there is either an abstract representation of the program which is modified in its realization by contextual variables, or that there is a sort of "go-to" instruction which provides access to a motor output program system. The latter is probably more likely, although some relatively specific information concerning, for example, deviant pronunciation of foreign words, may be included. With respect to the perceptual/identification component, it seems most useful to consider that as an address in memory specified by the perceptual features of the word.

The semantic component of the permanent representation of a word is assumed to include both semantic information and certain syntactic information. The grammatical (syntactic) hypothesis proposed by the present account follows the

Case Grammar theories of Fillmore (1968, 1971) and Anderson (1971). Very briefly, these theories assume that the syntactic analysis of a sentence assigns roles to the constituents of the sentence such that a given noun or noun phrase, for example, may be identified as the Agent of the sentence, another as the Object of the action or transformation specified by the verb, and so on. In terms of the present theory of lexical memory, it is assumed that the roles the word may play in sentences are specified in the semantic component of the representation, and that in conjunction with each role so specified there is semantic information specifying the meaning(s) that the word may take in that role. The semantic information itself is assumed to be represented as a list or pattern of features in a manner similar to that suggested by Smith, Shoben, and Rips (1974; it should be noted that the present theory follows Smith, et al., only with respect to structural assumptions, and does not use their processing hypotheses). The sort of grammatical analysis proposed is consistent with current "structural meaning" theories of language acquisition (e.g., Bloom, 1970; Schlesinger, 1971), which invoke roles like Agent, Object, and so on in their descriptions of child language. The proposal of several meanings, or senses, of a word being stored together follows Reder, Anderson, and Bjork (1974). That the three components of the representation of a word in permanent lexical memory are hypothesized to be semi-independent has support in data, for example the tip-of-the-tongue data of Brown and McNeill (1966), and is necessary to account for the varying effectiveness of different types of cues in verbal recall (see Tulving & Thomson, 1973, for a review).

Understanding a sentence, in this account, occurs when the person (a) assigns a grammatical role to each constituent; (b) accesses the permanent representation of each constituent word by means of the perceptual/identification information; (c) retrieves the semantic information stored under the role designation assigned; and (d) sorts from among the possible semantic interpretations of each constituent that which gives an unambiguous reading for the sentence as a whole. The last step, of course, requires coordinating the semantic information across constituents, and may be governed by some sort of selectional restrictions or rules.

The present theory assumes, consistent with current thought, (e.g., Craik & Tulving, 1975) that the memorial representation of an input stimulus event is the product of the processing performed on the stimulus. That processing is governed in part by the requirements of the task set for the person at input, and in part by information available in permanent memory. With specific reference to verbal materials, processing a word in a semantic task requires the use of semantic information stored at the permanent representation of the word in lexical memory. What is stored as the trace of that input event then consists of (a) the information from permanent lexical memory used in performing the input task; (b) information about the episodic context of the input stimulus; and (c) in the case that the word is part of a sentence, information concerning the word's grammatical (case) role, stored in conjunction with the semantic information. It is proposed that the traces of the content words (Fillmore, 1971) of a sentence are what are stored; function words are not stored as part of the memorial representation of a sentence, but rather are used by the grammar to achieve a semantic reading of the sentence, and are generated in the production of a sentence as necessary to yield the semantic content desired by the speaker. That is, function words are assumed to be part of the syntax of the sentence rather than part of its semantic structure proper, and while they do serve in part to determine the semantic reading given a sentence they do not

directly participate in that semantic reading or in the memorial trace of a sentence that has been processed as a semantic entity.

Tulving (1972; Tulving & Thomson, 1973; Watkins & Tulving, 1975) has developed an episodic theory of memory which asserts that the trace of an input event, e.g., a word, is unique in that it is not constructed as part of the pre-existing memorial lexicon or semantic memory system, but rather is separate from and independent of permanent semantic memory (Watkins & Tulving, 1975, p. 7). The present proposed theory rejects that assertion on both logical and empirical grounds. First, if each input word or sentence is stored as a separate independent trace, there is no means provided for the incorporation of new knowledge into existing knowledge structures, and it is also difficult to understand how such information is accessed by a retrieval mechanism. That is, if new episodic traces are independent of old knowledge, then the structure of old knowledge (the permanent lexical memory system, in this account) is useless as an organizational device for the recall process, and no clear alternative retrieval structure is posited in Tulving's work. Second, it is demonstrably the case that there are circumstances in which utilization of the organizational structure of permanent memory aids retrieval from episodic memory (e.g., Bilodeau & Fox, 1968; Fox & Dahl, 1971). It is equally demonstrable that there are conditions under which utilization of the organizational structure of the permanent lexical memory system interferes with recall (Hoppe & Dahl, Note 4; Tulving & Thomson, 1973) or is neutral with respect to recall (Hoppe & Wipper, Note 6), at least as compared with retrieval aids consisting of material explicitly presented to the person at the time of acquisition of the to-be-remembered words. As a consequence, the assumption that the trace of an input event (word or sentence) is independent of any information from permanent memory which might be used in constructing the trace (Watkins & Tulving, 1975) appears untenable, and it seems more in keeping with the various findings of interaction in retrieval between the permanent system and the episodic traces constructed during experimental acquisition of verbal materials to assume that the two systems, permanent lexical memory and episodic memory traces, share an organizational structure.

The present theory asserts that given an input sentence, e.g., The rock crushed the hut, the three content words (rock, crushed, hut) are assigned grammatical (case) roles by the processing mechanism. The permanent representation of each is interrogated, first to confirm that the word can take the case assigned by the processor, and then to ascertain the semantic meaning(s) each may take given that role. By coordinating the semantic meanings of the three elements, i.e., by selecting a meaning for each that is consistent with the meanings selected for the other two, a semantically unambiguous reading for the whole sentence is obtained: the person knows what the sentence meant; he or she understands the sentence. What is stored, however, is not the meaning of the whole sentence in a single 'clump', but rather the information used in the comprehension process, namely, information about the case roles played by the constituents, information respecting the final semantic reading given each, and information about the episodic context in which each constituent occurred, i.e., the grammatical and semantic information used in interpreting the other elements of the sentence. Each content word of the sentence has its information stored in organizational proximity to the representation in permanent lexical memory of the information used. The semantic information (case role) and episodic context used in understanding rock is stored in organizational proximity to the representation of rock in permanent lexical memory, and in particular, it is stored in organizational proximity to the specific semantic information in that permanent representation that was used in understanding the sentence.

Similarly, the semantic meaning, case role, and episodic surround of crushed are stored in organizational proximity to the representation of crush in lexical memory, and the same is the case with hut. (I should note that I have drastically over-simplified my presentation of the grammatical analysis assumed; crushed, for example, must be analyzed in terms of sense, tense, and other characteristics of verbs of transformation. Fillmore, 1971, discusses an analogous analysis for verbs of motion.) As a consequence, the sentence as it is stored shares the organizational structure of the permanent lexical memory system because the elements of the sentence are stored in organizational proximity to their respective representations in that permanent system. In addition, however, the organizational structure of the sentence itself is also preserved, since information concerning the case role (and hence by inference the necessity for other case roles played by other elements) and information concerning the episodic context in which each element occurred, the latter consisting in part of information about the other elements of the sentence, is stored in the trace of each element. There are, then, two organizational structures superimposed upon one another, one for permanent lexical memory, and the other for episodic memories. The two are not conceived of as being organized by links connecting locations (e.g., by associative bonds), but rather the organization is implicit in the information stored at the various organizational locations, and is realized in the utilization of the information in the processing of information. One might even speculate that the permanent lexical memory system itself has at least three organizational schemes intermingled; i.e., the location of the representation of a given word might well be coded in terms of perceptual/identification information, response generation information, and at least one semantic information system, and the same item may be locatable in at least three multi-dimensional organizational spaces. Development of that speculation is, however, beyond the scope of this report.

Given the foregoing sketch for theoretical background, an example may be useful in clarifying the memorial structures postulated to be formed given sentences as input. Suppose the person in question hears the sentence The rock rolled down the mountain. The episodic structure formed will, it is hypothesized, look like that depicted in Figure 9a. Note that the constituent

---

Insert Figure 9 about here

---

down the mountain is represented in the figure as a single entity. This is purely for expository convenience; in fact it is assumed that the phrase would be analyzed into two constituents, a Locative (mountain) and a Directional (down), in keeping with the multi-argument nature of the verb of motion rolled. The episodic representation of each constituent contains (a) the case role played by the constituent, (b) the semantic information from the permanent representation of the constituent that was used in understanding the sentence, abbreviated in the figure S.I., and (c) a complex of episodic information (abbreviated E.I.) which contains spatio-temporal information about the occurrence of the item and information about other constituents. This latter information, about other constituents, is not a representation of the other words of the sentence, but is rather 'post-processing' information, including information concerning other case roles present in the sentence and semantic information about them.

Suppose, now that a second sentence occurs: The rock crushed the hut. In order for the rock of this sentence to be identified with the rock of the first sentence, the processing mechanism must be able to access the representation constructed at the time of occurrence of the first sentence. That access is by means of the permanent representation in lexical memory. The Tulving account of episodic memory, asserting as it does that the episodic representation of the first sentence is independent of permanent memory, provides no means for the identification of the second rock with the first, and therefore provided no account of the well replicated finding (Bransford & Franks, 1971; Cofer, 1973; Curnow, 1969; Hoppe, Note 2, Note 3; Katz, 1973) that subjects do appear to identify the second occurrence of rock with the first, and construct a representation of the whole situation specified by the composite sentence The rock rolled down the mountain and crushed the hut. The present theory asserts that the person does identify the two as indicating the same Agent, and produces the representation shown in Figure 9b. In both Figure 9a and Figure 9b the episodic trace of a sentence element is formed in organizational proximity to the semantic component of the representation of the word in permanent lexical memory, and in particular in organizational proximity to the specific semantic information used in constructing the trace, and this is the means of access to the episodic trace. The permanent memory representations are left off the figures to avoid overcomplexity of the drawing.

Contrast the structure of Figure 9b with that constructed given the sentence (to a naive person) The rock rolled down the mountain and crushed the hut at the edge of the woods, shown in Figure 10. In this structure, there is no differentiation in the episodic information among the sentence elements. Under these circumstances, people should be more likely to 'recognize' The rock rolled down the mountain and crushed the hut in a test than in the circumstances yielding the structure of Figure 9b. This outcome is predicted on the basis of the discriminability of the episodic information contained in the two structures: With reference to Figure 9b, the test sentence contains elements with different episodic information, while with reference to Figure 10, the episodic information is the same for the elements of the test sentence.

The general form of the theory sketched above is not unique; Anderson and Bower (1973) and Norman and Rumelhart (1975, esp. Chapters 1 & 8) have suggested elaborate theoretical schemes, simulatable on computers, which share some characteristics with the above account. Indeed, certain aspects of those accounts may be incorporated into the present theory, with some modification. The present suggestion differs from those cited in several important respects, however. Most important, it offers a way of accounting for the fact that subjects do see the relationships among the acquisition sentences from the various idea sets, segregate them into separate representations, construct representations such that inferences about new but consistent test sentences may be made, and still discriminate old from new sentences. This last is what the other accounts are incapable of doing. In Norman and Rumelhart's (1975) theory, for example, once a second acquisition sentence is incorporated into the structure in memory, there is no information whatsoever on the basis of which to say which components of the representation came when. All episodic information is lost, in that it is not represented in the memory structure constructed.

Second, the present suggestion explicitly includes a description of permanent lexical memory which is congruent with current processing views of word memory (e.g., Craik & Tulving, 1975). That is, the present account offers a potential means of subsuming both the literature on memory for word lists and the work on sentence memory presented (and cited) in this report. Howes (Note 5) has

shown that the same sorts of processing variables that Craik and Tulving (1975) identify as dominant in word memory also strongly affect sentence memory in the Bransford-Franks experimental procedure. Her work was motivated by theoretical considerations that were early forms of the account outlined above.

Reitman and Bower (1973) offered an account of the Bransford-Franks phenomena that is independent of semantic considerations. They proposed a "tally model" in which the memory simply tabulates the frequency of occurrence of elements of the complex stimuli, and makes the recognition judgments on the basis of the total tallies for a given test sentence. Reynolds and Flagg (1977) criticize Reitman and Bower's formulation on the grounds that Reitman and Bower used non-semantic materials in their experiments, and as a consequence their results do not necessarily speak to semantic memory. In addition, a simple tally model cannot account for the results reported here; the new and old test sentences in the present experiments did not differ with respect to the frequency of occurrence of the elements in acquisition, and as a consequence the subjects had no reason on that basis to make a discrimination. Again, what is necessary is an account which can simultaneously predict discrimination of old from new test sentences and predict the slope of the recognition curve with respect to test sentence inclusiveness.

To summarize the theoretical discussion: Tulving's distinction between episodic and semantic memory, while possibly heuristically useful in the domain of word memory, does not allow for the integration of nonconsecutively experienced sentences into structures that can underlie people's mistaken recognition of new but consistent sentences in test in the Bransford-Franks procedure. On the other hand, the theories of Anderson and Bower (1973) and Norman and Rumelhart (1975), while allowing such integration to occur, do not have a means to account for subjects' ability to still discriminate new from old information, even while they are showing the curve for new sentences that is taken to indicate integration. The same is true of Reitman and Bower's (1973) account. What is necessary is an account which allows both to happen. That is what is suggested in this report.

The results also have implications for educational practice. When the material to be learned in a classroom or text situation consists of an interrelated set of concepts, and the object of the learning experience is to bring the learner to an understanding of both the specific materials presented and potential relationships not explicitly presented, then the results suggest that the various subsets of the total materials be presented several times, in several different forms. With only one presentation, as in the 1/3 condition of Experiment 2, people are not drawing the inferences from the material that are possible. Only after three presentations of each semantic relation do people appear to be bringing the various components together into a wholistic representation. In other words, inferential memory depends, it appears, upon repetition of the component materials. This echoes the traditional associationists' emphasis upon repetition, though for different, meaningful materials. The research also suggests that the inferential memory that is constructed over a full acquisition experience is relatively enduring; the research participants' willingness to accept inferred relationships was consistent over a one week retention interval.

Reference Notes

1. Gould, D.R., & Bourland, A.M. Recognition memory for sentences as a stimulus discrimination task. Paper presented at the Northwestern Pennsylvania Psychological Association Undergraduate Conference, Cambridge Springs, Pennsylvania, April, 1973.

2. Hoppe, R.B. Semiannual Report on National Institute of Education Project No. 4-0404, December, 1974.

3. Hoppe, R.B. Semiannual Report on National Institute of Education Project No. 4-0404, June, 1975.

4. Hoppe, R.B., & Dahl, P.R. Prompted recall in serial learning. Manuscript submitted for publication, 1977.

5. Howes, C.A. The effects of depth of processing on integration of semantically related sentences. Unpublished manuscript, 1975. (Available in Kenyon College Library, Gambier, OH 43022).

6. Hoppe, R.B., & Wipper, D.P. The effect of cue type on retrieval with a semantic orienting task. Unpublished manuscript, 1977. (available from R.B. Hoppe, Kenyon College, Gambier, OH 43022).

References

Anderson, J.M. 1971. The grammar of case: towards a localistic theory. Cambridge: Cambridge University Press.

Anderson, J.R., & Bower, G.H. Human associative memory. Washington, D.C.: Winston.

Barclay, J.R. 1973. The role of comprehension in remembering sentences. Cognitive Psychology, 4, 229-254.

Bilodeau, E.A., & Fox, P.W. 1968. Free association, free recall, and stimulated recall compared. Behavior Research Methods and Instrumentation, 1, 14-17.

Bloom, L. 1970. Language development: form and function in emerging grammars. Cambridge: MIT Press.

Bransford, J.D., & Franks, J.J. 1971. The abstraction of linguistic ideas. Cognitive Psychology, 2, 331-350.

Bransford, J.D., Barclay, J.R., & Franks, J.J. 1972. Sentence memory: a constructive versus interpretive approach. Cognitive Psychology, 3, 193-209.

Brown, R.W., & McNeill, D. 1966. The tip-of-the-tongue phenomenon. Journal of Verbal Learning and Verbal Behavior, 5, 325-337.

Cofer, C.N. 1973. Constructive processes in memory. American Scientist, 61(5), 537-543.

Craik, F.I.M., & Tulving, E. 1975. Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 104, 268-294.

Curnow, P.F. 1969. Integration of linguistic materials. Unpublished Ph.D. dissertation, University of Minnesota.

Ebbinghaus, H. 1885. Über das Gedächtnis: Untersuchungen zur experimentelen Psychologie. Leipzig: Duncker and Humbolt.

Fillmore, C.J. 1968. The case for case. In Bach, E., & Harms, R.T. (eds.), Universals in linguistic theory. New York: Holt, Rinehart and Winston.

Fillmore, C.J. 1971. Types of lexical information. In Steinberg, D.D., & Jakobovits, L.A. (eds.), Semantics. Cambridge: Cambridge University Press.

Flagg, P.W. 1976. Semantic integration in sentence memory? Journal of Verbal Learning and Verbal Behavior, 15, 491-504.

Fox, P.W., & Dahl, P.R. 1971. Aided retrieval of previously unrecalled information. Journal of Experimental Psychology, 88, 349-353.

Franks, J.J. & Bransford, J.D. 1971. Abstraction of visual patterns. Journal of Experimental Psychology, 90, 65-74.

Franks, J.J., & Bransford, J.D. 1972. The acquisition of abstract ideas. Journal of Verbal Learning and Verbal Behavior, 11, 311-335.

Gibson, E.J. 1971. Perceptual learning and the theory of word perception. Cognitive Psychology, 2, 351-368.

Jarvella, R.J. 1971. Syntactic processing of connected speech. Journal of Verbal Learning and Verbal Behavior, 10, 409-416.

Katz, S. 1973. Role of instructions in abstraction of linguistic ideas. Journal of Experimental Psychology, 98, 79-84.

Norman, D.A., & Rumelhart, D.E. 1975. Explorations in cognition. San Francisco: Freeman.

Reder, L.M., Anderson, J.R., & Bjork, R.A. 1974. A semantic interpretation of encoding specificity. Journal of Experimental Psychology, 102, 648-656.

Reitman, J.S., & Bower, G.H. 1973. Storage and later recognition of exemplars of concepts. Cognitive Psychology, 4, 194-206.

Reynolds, A.G., & Flagg, P.W. 1977. Cognitive psychology. Cambridge: Winthrop.

Sachs, J.S. 1967. Recognition memory for syntactic and semantic aspects of connected discourse. Perception and Psychophysics, 2(9), 437-442.

Schlesinger, I.M. 1971. Learning grammar: from pivot to realization rule. In Huxley, R., & Ingram, E. (eds.) Language acquisition: models and methods. New York: Academic Press.

Smith, E.E., Shoben, E.J., & Rips, L.J. 1974. Structure and process in semantic memory: a featural model for semantic decisions. Psychological Review, 81, 214-241.

Strange, W., Keeney, T., Kessel, F., & Jenkins, J.J. 1970. Abstraction over time from distortions of random dot patterns. Journal of Experimental Psychology, 83, 508-510.

Tulving, E. 1972. Episodic and semantic memory. In Tulving, E., & Donaldson, W. (eds.), Organization of memory. New York: Academic Press.

Tulving, E., & Thomson, D.M. 1973. Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80, 352-373.

Vaughan, G.M., & Corballis, M.C. 1969. Beyond tests of significance: estimating strength of effects in selected ANOVA designs. Psychological Bulletin, 72, 204-213.

Watkins, M.J., & Tulving, E. 1975. Episodic memory: when recognition fails. Journal of Experimental Psychological: General, 104, 5-29.

Wickens, D.D. 1970. Encoding categories of words: an empirical approach to meaning. Psychological Review, 77, 1-15.

Table 1

Median Oddity Ratings for FOURs and NONCASEs:

Experiment 1

Condition	Oddity Rating	
	FOURs	NONCASEs
Normal-Odd	2.00	4.25
Odd-Normal	3.75	1.67
Normal-Normal	2.41	2.95
Odd-Odd	3.83	3.90

Table 2

Mean Recognition Scores for Old Test Sentences: Experiment 2.

Condition	Sentence Type		
	THREE	TWO	ONE
1/3	+3.00	+2.64	+0.33
2/3	+3.17	+2.50	+0.08
3/3	+2.57	+2.14	+1.93

Table 3

Analyses of Variance: Experiment 2

Group	Source	df	MS	F
1/3	Old/New (A)	1	203.68	19.32**
	Error (A x S)	11	10.54	
	Sentence Type (B)	2	24.77	3.01
	Error (B x S)	22	8.24	
	A x B	2	5.22	0.65
	Error (A x B x S)	22	8.05	
2/3	Old/New (A)	1	196.35	53.07***
	Error (A x S)	11	3.70	
	Sentence Type (B)	2	67.90	11.30***
	Error (B x S)	22	6.01	
	A x B	2	0.92	0.19
	Error (A x B x S)	22	4.97	
3/3	Old/New (A)	1	45.28	7.78**
	Error (A x S)	11	5.82	
	Sentence Type (B)	2	33.56	7.17**
	Error (B x S)	22	4.68	
	A x B	2	18.52	4.68*
	Error (A x B x S)	22	3.96	

*$p < .05$

**$p < .01$

***$p < .001$

Table 4

Analyses of Variance: Experiment 3

Group	Source	df	MS	F
1/3	Overlap (A)	1	0.75	0.13
	Error (A x S)	11	15.06	
	Sentence Type (B)	3	64.98	6.87**
	Error (B x S)	33	9.46	
	A x B	3	19.62	1.73
	Error (A x B x S)	33	11.33	
2/3	Overlap (A)	1	53.13	6.26*
	Error (A x S)	11	8.48	
	Sentence Type (B)	3	143.18	20.48***
	Error (B x S)	33	6.99	
	A x B	3	29.74	6.42**
	Error (A x B x S)	33	4.63	
3/3	Overlap (A)	1	7.97	4.80*
	Error (A x S)	11	1.66	
	Sentence Type (B)	3	115.09	20.02***
	Error (B x S)	33	5.75	
	A x B	3	7.68	1.13
	Error (A x B x S)	33	6.77	

*$p < .05$

**$p < .01$

***$p < .001$

Table 5

Mean Recognition Scores for Old Sentences:

Experiment 4

Condition	Sentence Type		
	THREE	TWO	ONE
1-Hour	+4.83	+2.42	+1.04
1-Day	+4.92	+2.10	+1.67
1-Week	+3.92	+2.04	+0.96

Table 6

Analyses of Variance: Experiment 4

Group	Source	df	MS	F
1-Hour	Old/New (A)	1	114.76	38.90***
	Error (A x S)	11	2.95	
	Sentence Type (B)	2	94.03	50.28***
	Error (B x S)	22	1.87	
	A x B	2	2.59	0.61
	Error (A x B x S)	22	4.22	
1-Day	Old/New (A)	1	47.69	10.08**
	Error (A x S)	11	4.73	
	Sentence Type (B)	2	104.88	32.17***
	Error (B x S)	22	3.26	
	A x B	2	14.68	8.90**
	Error (A x B x S)	22	1.65	
1-Week	Old/New (A)	1	52.87	9.72**
	Error (A x S)	11	5.44	
	Sentence Type (B)	2	106.93	18.06***
	Error (B x S)	22	5.92	
	A x B	2	11.08	4.03*
	Error	22	2.75	

*$p < .05$

**$p < .01$

***$p < .001$

Table 7

Estimates of Variance Accounted For:

Experiments 2 & 4

Experiment	Condition	Old/New	Sentence Type	Interaction	Total	Pooled Error
2	1/3	23.2	0.0	0.0	23.2	8.942
	2/3	28.6	18.8	0.0	47.4	4.891
	3/3	8.6	12.2	5.8	26.6	4.820
4	1-Hour	21.9	35.6	0 0	57.5	3 012
	1-Day	8.9	40.5	4.6	54.0	3.210
	1-Week	8.0	33.9	2.1	44.0	4.702

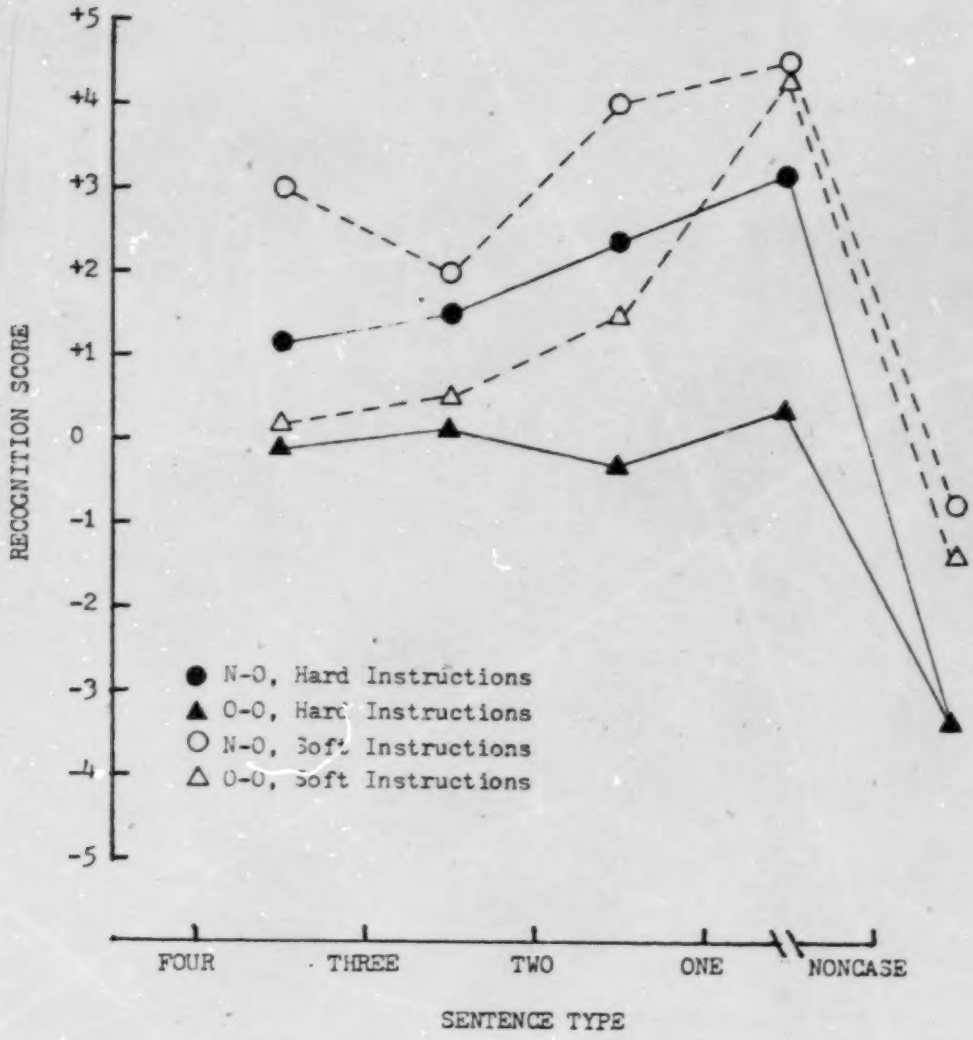

Figure 1. 'Recognition' scores for N-O and O-O materials in the two No-Acquisition conditions of Experiment 1.

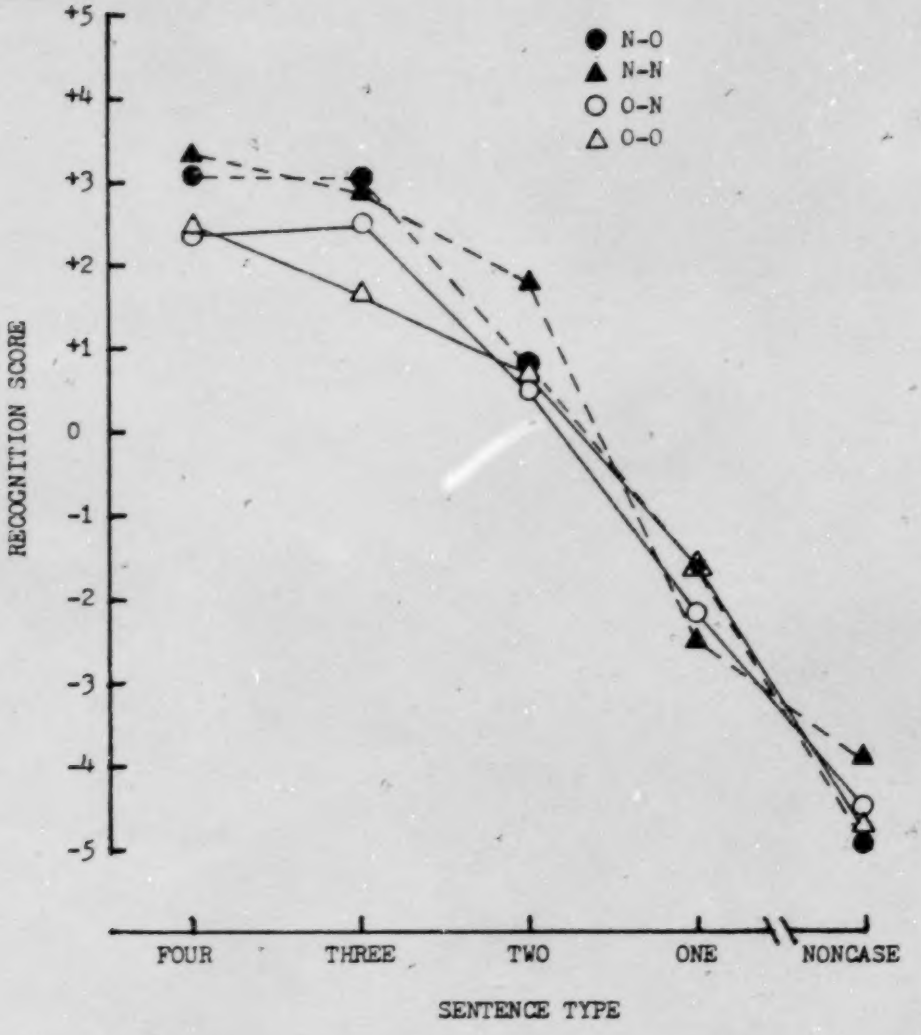

Figure 2. Recognition scores averaged over subjects for four materials conditions of Experiment 1.

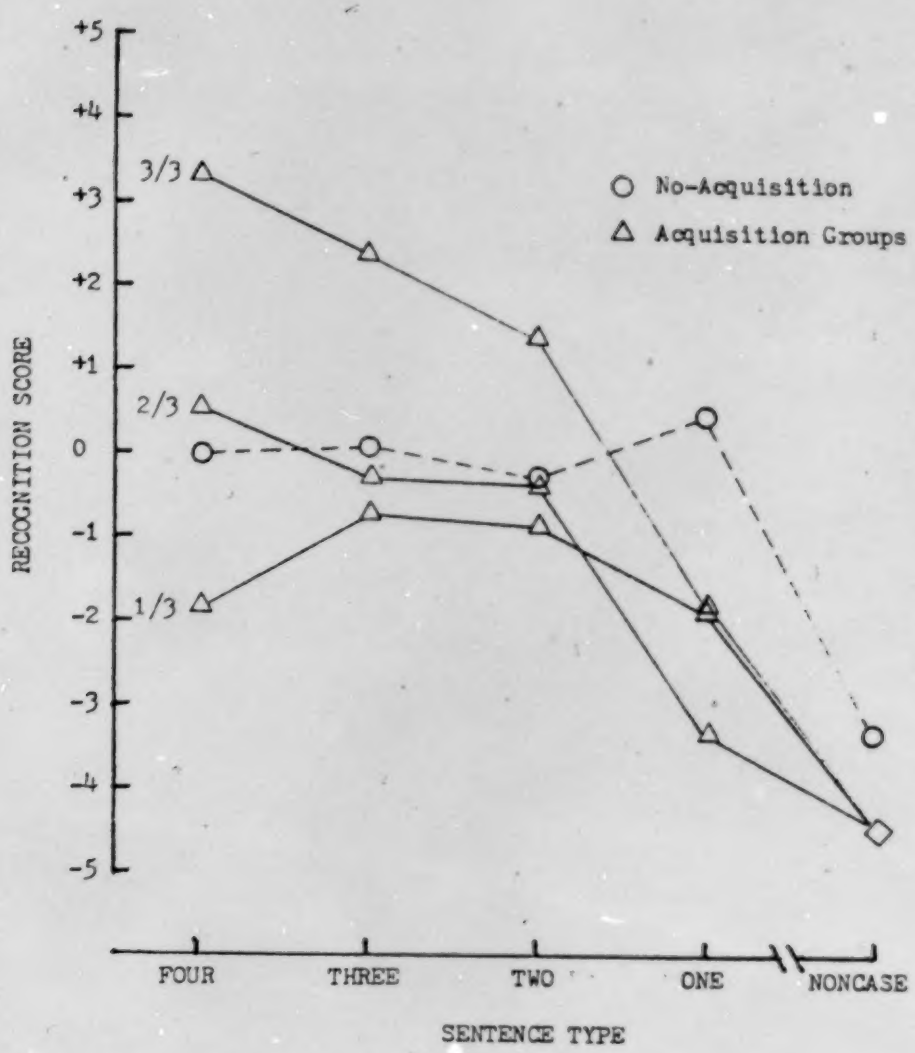

Figure 3. Mean recognition scores for New test sentences in Experiment 2.

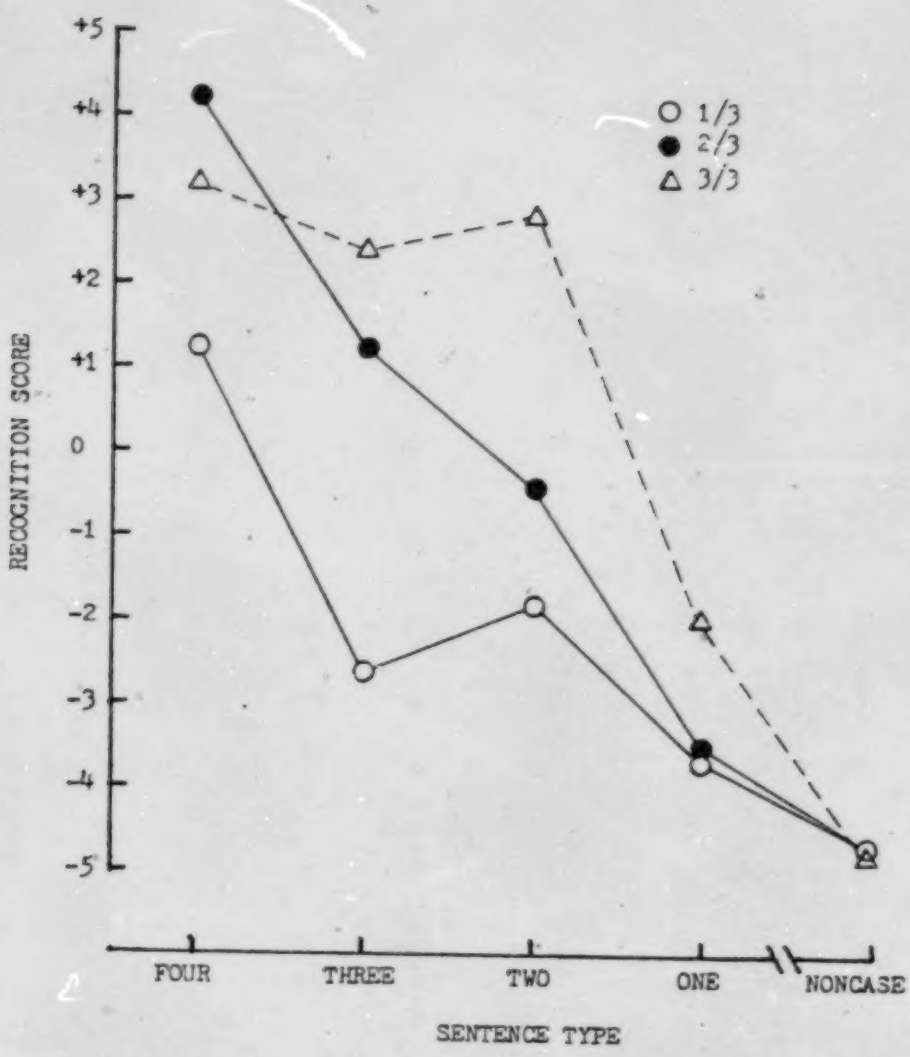

Figure 4. Recognition scores for New test sentences of the Overlap Idea Sets in three acquisition conditions of Experiment 3.

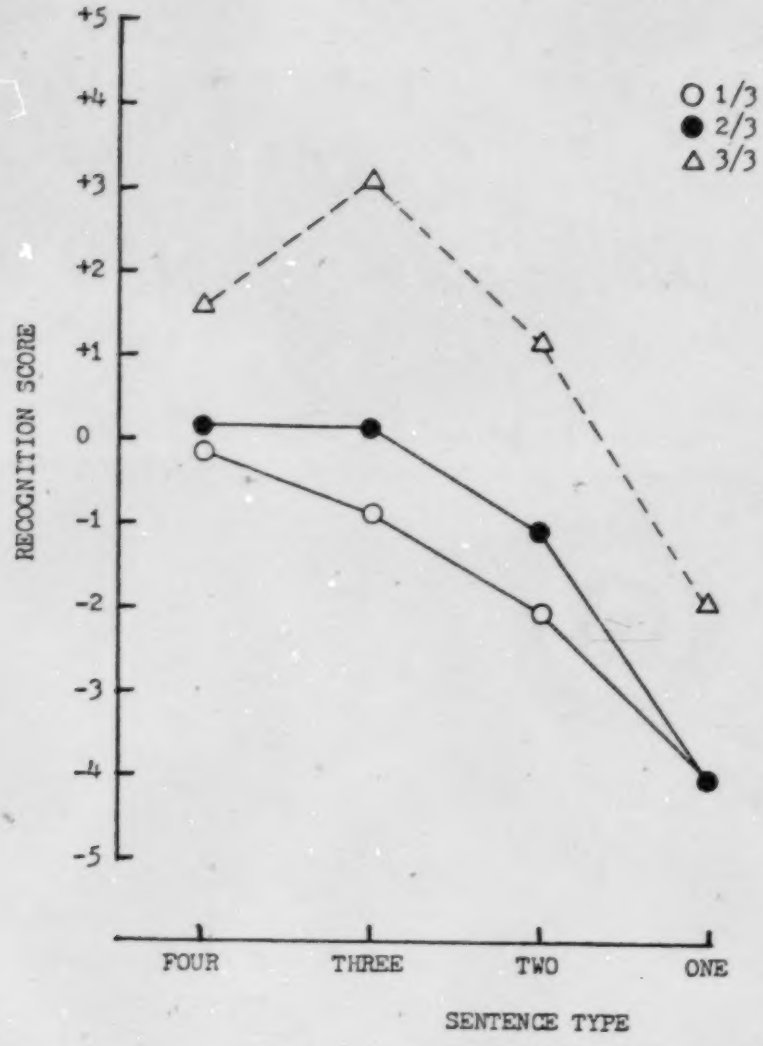

Figure 5. Recognition scores for New test sentences of the Nonoverlap Idea Sets in three acquisition conditions of Experiment 3.

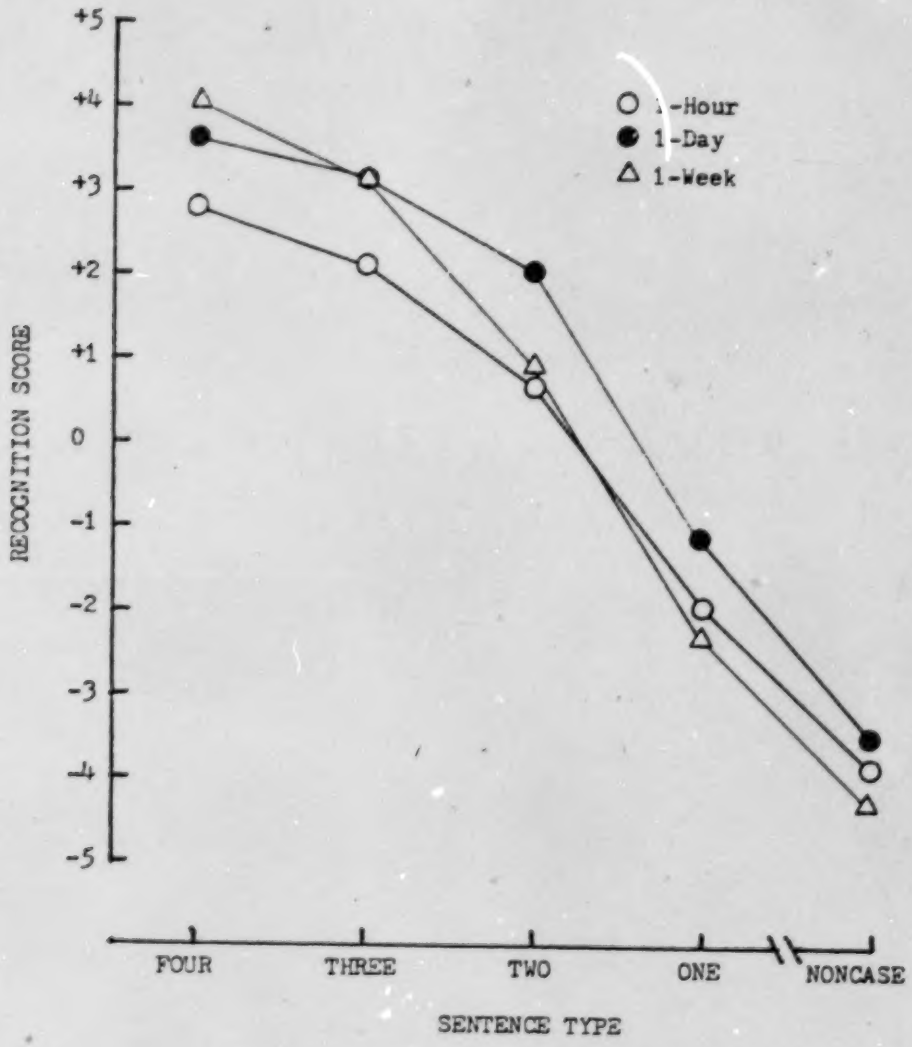

Figure 6. Recognition scores for New test sentences in three retention conditions of Experiment 4.

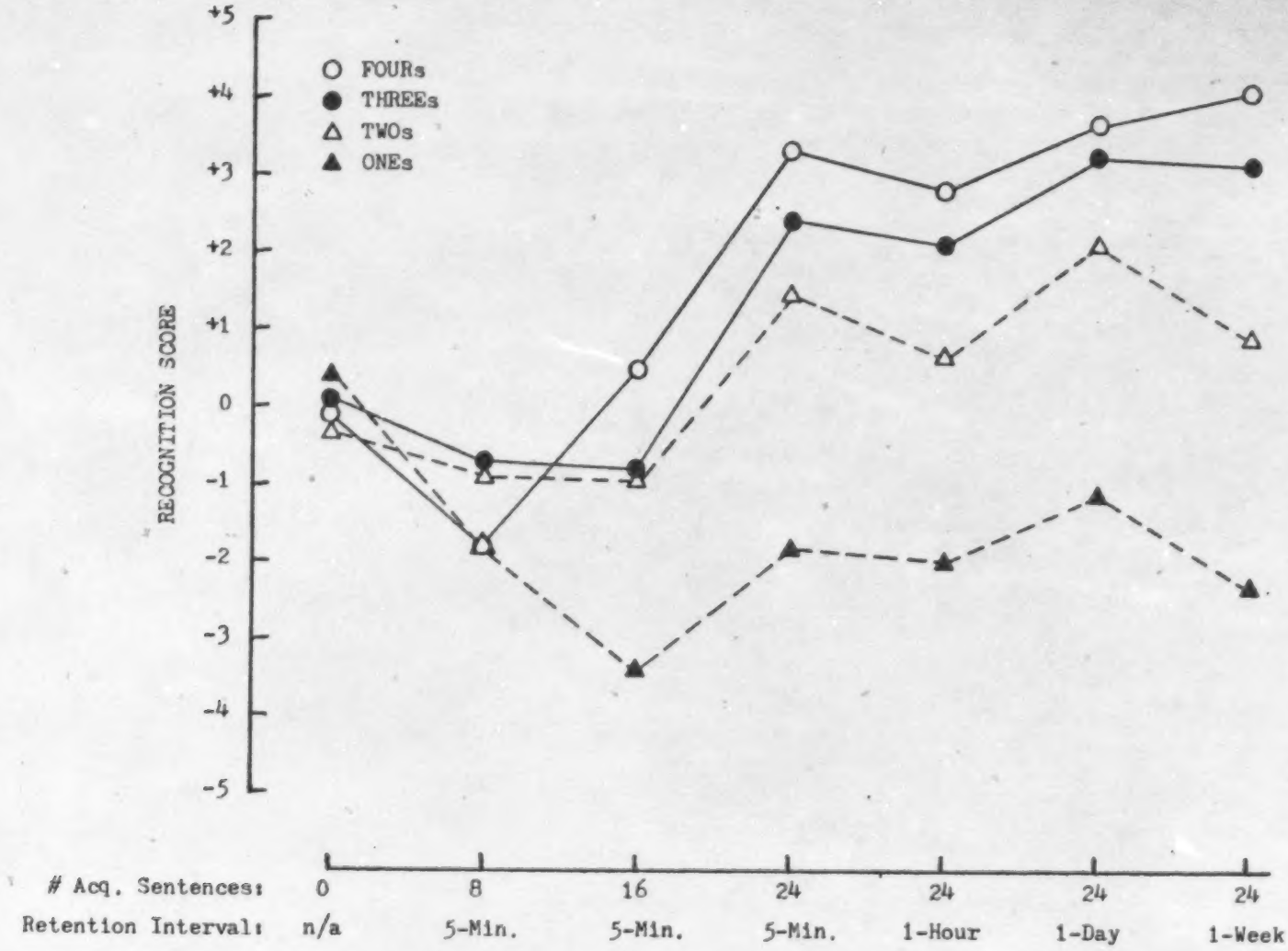

Figure 7. Recognition scores for New test sentences in the various acquisition and retention conditions of Experiments 2 and 4.

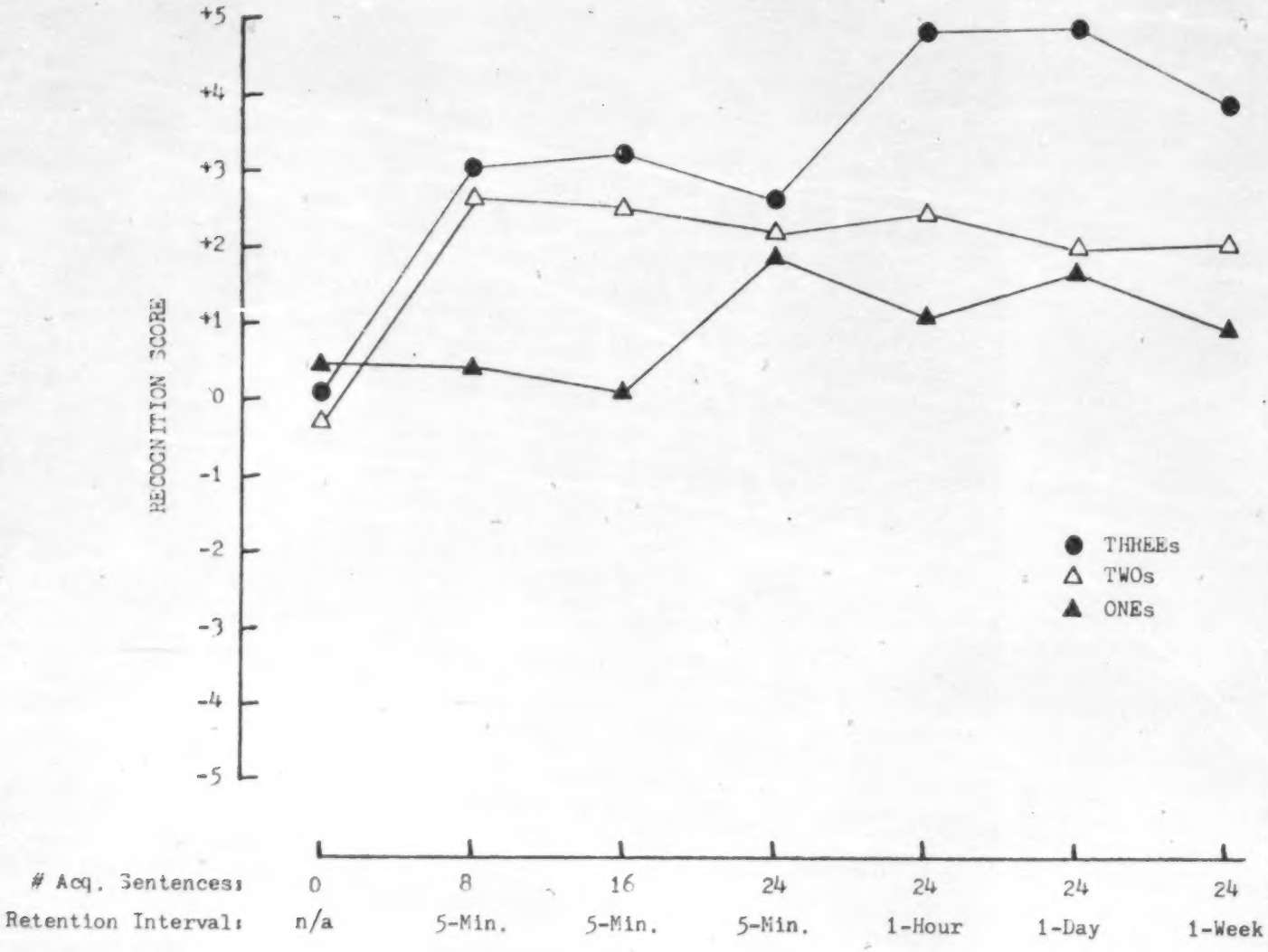

Figure 8. Recognition scores for Old test sentences in the various acquisition a retention conditions of Experiments 2 and 4.

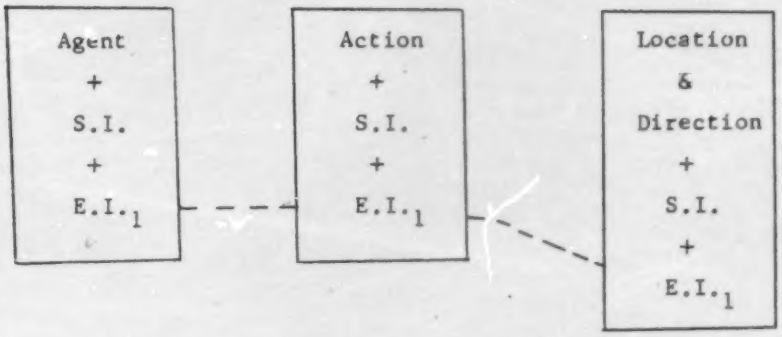

Figure 9a. The configuration of the episodic trace produced by processing the sentence The rock rolled down the mountain.

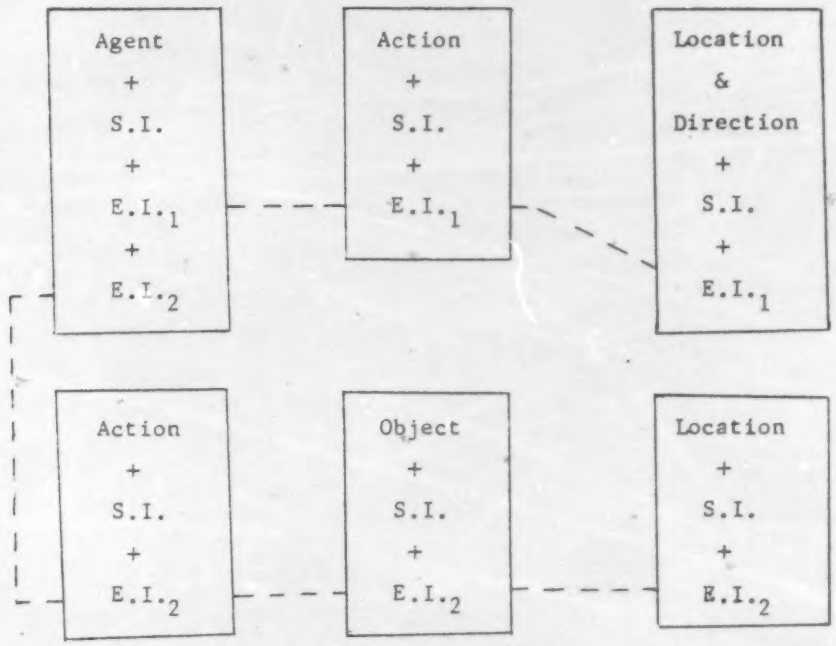

Figure 9b. The configuration of the episodic trace produced by the addition of another sentence, The rock crushed the hut at the edge of the woods, to the sentence of Figure 1a.

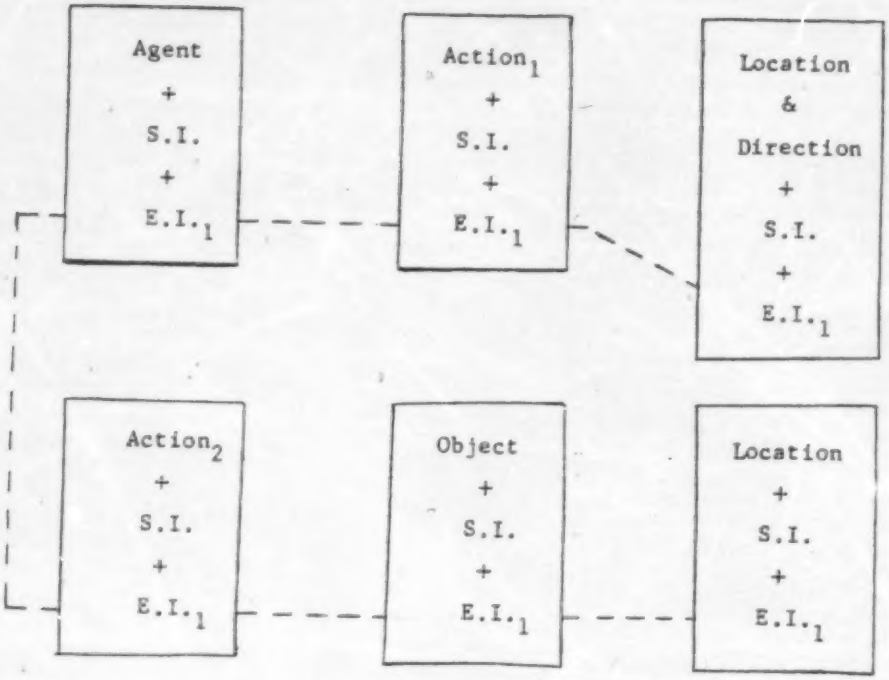

Figure 10. The configuration of the episodic trace produced by processing the sentence The rock rolled down the mountain and crushed the hut at the edge of the woods.